

Which Models of Scientific Explanation are (In)Compatible with IBE?

Yunus Prasetya

Forthcoming in *The British Journal for the Philosophy of Science*

(This is the penultimate draft, prior to copyediting)

<https://doi.org/10.1086/715203>

Abstract

In this article, I explore the compatibility of inference to the best explanation (IBE) with several influential models and accounts of scientific explanation. First, I explore the different conceptions of IBE and limit my discussion to two: the heuristic conception and the objective Bayesian conception. Next, I discuss five models of scientific explanation with regard to each model's compatibility with IBE. I argue that Philip Kitcher's unificationist account supports IBE; Peter Railton's deductive-nomological-probabilistic model, Wesley Salmon's statistical-relevance Model, and Bas van Fraassen's erotetic account are incompatible with IBE; and Wesley Salmon's causal-mechanical model is merely consistent with IBE. In short, many influential models of scientific explanation do not support IBE. I end by outlining three possible conclusions to draw: (1) either philosophers of science or defenders of IBE have seriously misconstrued the concept of explanation, (2) philosophers of science and defenders of IBE do not use the term 'explanation' univocally, and (3) the ampliative conception of IBE, which is compatible with any model of scientific explanation, deserves a closer look.

1. *Introduction*
2. *What is IBE?*
3. *An Account that Supports IBE: Kitcher's Unificationist Account*
4. *Models that are Incompatible with IBE*
 - 4.1. *Peter Railton's deductive-nomological-probabilistic (DNP) model*
 - 4.2. *Wesley Salmon's statistical-relevance (SR) model*
 - 4.3. *Bas van Fraassen's erotetic account*
5. *A Model that is (Merely) Compatible with IBE: Salmon's Causal-Mechanical (CM) Model*
6. *Conclusion*

1 Introduction

Explanationists argue that inference to the best explanation (IBE) is a rational form of inference. However, they are rarely concerned with how IBE relates to models and accounts of explanation discussed in philosophy of science. They tend to focus only on theoretical virtues—simplicity, consilience, mathematical elegance, and so on—that might be relevant to confirmation. Likewise, philosophers who discuss models of scientific explanation typically are not concerned with how or whether a particular model can support IBE.

Why does this matter? For one, explanationists often claim that IBE is ubiquitous in ordinary life and essential for scientific reasoning. See, for example, (Thagard [1978]; Boyd [1983]; McMullin [1992]; Okasha [2000]; McGrew [2003]; Lipton [2004]; Douven [2013]; McCain and Poston [2014], [2018], [2019]; McCain [2015]; Bird [2018]; Cabrera [2020]; Schupbach [2018]). Therefore, the philosophical discussion on the nature of scientific explanation seems relevant to explanationism. Models of scientific explanation purport to tell us about the nature of explanation in science. However, some of these models are incompatible with IBE, so explanationists may make enemies out of the philosophers who defend these models. Additionally, there are models of scientific explanation that support IBE. Explanationists may find allies in those philosophers who defend IBE-friendly models of scientific explanation.

Second, some philosophers affirm the need for an investigation on the intersection between IBE and models of scientific explanation. Wesley Salmon ([1989]) and Peter Lipton ([2001]), though they disagree on whether IBE is rational, agree that IBE requires a substantive account of explanation.¹ Frank Cabrera ([2020]) discusses the plentitude problem: the objection that IBE is indeterminate because there are plenty of competing models of scientific explanation, each with its own merits and demerits.²

In this paper, I respond to the need for a discussion about IBE and models of scientific explanation. I explore some of the most influential models of scientific explanation, especially regarding their compatibility with IBE. Specifically, I argue that Philip Kitcher's ([1976]) unificationist account supports IBE. On the other hand, Salmon's ([1971]) statistical-relevance model, Peter Railton's ([1978]) deductive-nomological-probabilistic model, and Bas van Fraassen's ([1980]) erotetic account are incompatible with explanationism. Finally, Salmon's ([1984]) causal-mechanical model, at best, is neutral with regard to explanationism—it does not support IBE, but it is compatible with IBE. These models and accounts of scientific explanation are among the most influential in the field. Each is independently plausible and most of them are featured in the *Stanford Encyclopedia of Philosophy* entry on scientific explanation by James Woodward ([2019]).

¹ See Schupbach and Sprenger ([2011]) for an argument to the contrary: we don't need an analysis of explanation to give an analysis of explanatory power that can support IBE.

² Cabrera ([2020]) outlines several possible responses to the plentitude problem that do not require investigating the different models of scientific explanation. Even if models of scientific explanation generally don't support IBE, the explanationist can defend IBE in some other way.

The exception, Railton's DNP model, was praised by Salmon for being 'quite possibly the best thing written on scientific explanation since Hempel's "Aspects" essay' ([1989], p. 120).

If I'm right, many models of scientific explanation are incompatible with IBE. So, philosophers of science and explanationists may have a fruitful discussion on explanation. Maybe one group has seriously misconstrued explanation and how practicing scientists actually use the concept of explanation. Alternatively, perhaps philosophers of science and explanationists do not use the term 'explanation' univocally.

The explanationist might object: Why can't we give an account of IBE that does not assume any particular model or account of scientific explanation? Schupbach and Sprenger take this approach in giving an account of explanatory power. They claim to provide 'several, more primitive adequacy conditions that [...] an analysis of explanatory power should satisfy' ([2011], 107).

Explanationists may use this kind of approach. However, they may end up with adequacy conditions that are incompatible with some of the most influential models of scientific explanation. Those already committed to one of these models, therefore, have a reason to reject explanationism. Similar points, of course, apply to those who develop and defend a particular model of scientific explanation. They should be mindful of whether their model is incompatible with IBE.

In the next section, I begin by exploring different conceptions of IBE. My argument will affect only two conceptions of IBE: the heuristic conception and the objective Bayesian conception. In section three, I discuss Kitcher's ([1989]) unificationist account and argue that it supports IBE. In section four, I explore Railton's DNP model, Salmon's SR model, and van Fraassen's erotetic account. I argue that they are all incompatible with IBE. In section five, I discuss Salmon's CM model and argue that it is compatible with IBE, but it does not support IBE. I conclude the paper in section six.

2 What is IBE?

We start our discussion on IBE by outlining several distinctions. First, explanatory reasoning is distinct from IBE. The former seems undeniably ubiquitous in science and ordinary life. We make judgments about whether a hypothesis explains a body of evidence, we compare different hypotheses, and judge some to be better explanations than others. Even critics of IBE can affirm that we often reason about explanations (van Fraassen [1980]; Salmon [2001a], [2001b]). Indeed, this is what a model of scientific explanation does; it gives us a set of criteria for what counts as an explanation or for evaluating explanations. However, one may ascertain that a hypothesis is the best explanation of a body of evidence and resist inferring anything about the truth value or probability of that hypothesis. So, IBE cannot be identified with explanatory reasoning. At most, IBE should be considered as a species of explanatory reasoning. IBE is a pattern

of inference starting from the explanatory features of a hypothesis to the truth value or likelihood of that hypothesis.

We should also distinguish between actual and potential explanations. The statement, ‘H explains E’, is often understood to analytically entail H and E. But this can’t be how IBE works. Otherwise, IBE would be no more useful than the inference from ‘S knows that P’ to ‘P is true’. To avoid trivializing IBE in this manner, we distinguish between potential and actual explanations. If H potentially explains E, then H fulfills all the criteria for being an explanation of E with the possible exception of truth.

Explanationists generally characterize IBE as a rule of inference that starts from the following premises (Harman [1965]; Psillos [2007]; Douven [2013]; McCain and Poston [2019]).

- 1) E is a body of evidence.
- 2) Hypothesis H potentially explains E better than any other hypothesis.
- 3) H is a sufficiently good potential explanation of E given our background information, K.³

What should we infer from these premises? Some explanationists suggest inferring that H is true (Harman [1965]; Psillos [2007]; McCain and Poston [2019]). Others suggest inferring that we have good reason to believe H or that we are justified in believing H (Peirce [1935]; Lycan [1988]; van Fraassen [1989]; Cabrera [2017]).

Alternatively, we may favor a more fine-grained approach by describing the conclusion probabilistically. I will outline the three most widely held conceptions of IBE in the literature and how they differ on the conclusion of IBE.⁴

The first is what we will call ‘the ampliative conception’, on which IBE is a distinctive and ampliative rule of inference based on the virtue of explanatory power. The ampliative conception is discussed in (Harman [1965]; Psillos [2007]; van Fraassen [1989]; Douven [2013]; Douven and Schupbach [2015]). Generally, defenders of the ampliative conception suggest using IBE after Bayesian conditionalization. Thus construed, IBE grants probabilistic bonuses to explanatory hypotheses after conditionalization on the evidence. Upon learning E, ampliative explanationists conclude that the posterior probability of H is greater than $P(H|E)$.

Second, we have the heuristic conception, wherein IBE plays a heuristic role for ordinary epistemic agents to approximate Bayesian reasoning. This conception is

³ There are exceptions to the generality. For example, on C. S. Peirce's ([1935]) classic formulation, E is noted to be surprising. Additionally, Douven and Mirabile ([2018]) provide empirical evidence that the average person infers the best explanation only when it is significantly better than its competitors. If we want to accommodate this finding, (2) should state that H is the best explanatory hypothesis and sufficiently better than its competitors.

⁴ There are, of course, other accounts that have been proposed recently. For example, Leah Henderson ([2014]) suggests an account on which IBE emerges from the practices of a Bayesian who adopts constraints on prior probabilities.

defended in (McGrew [2003]; Lipton [2004]; Dellsen [2018]). The conclusion of IBE, on the heuristic conception, is a statement that approximates the posterior probability of H as being equal or close to $P(H|E)$. For example, the heuristic explanationist might say that, upon learning E , the posterior probability of H is $r \pm \theta$, which is our best approximation of $P(H|E)$.

Finally, we have the objective Bayesian conception of IBE. On this conception, IBE fits within an objective Bayesian framework. IBE provides normative constraints on the prior probabilities that we use in Bayesian conditionalization. This conception is defended in (Weisberg [2009]; Poston [2014]; Climenhaga [2017]). In other words, the goodness of H , as a potential explanation of E , provides normative constraints for $P(H)$, or $P(E|H)$, or $P(E)$.

Defenders of the objective Bayesian conception do not always agree on how H 's status as a good explanation provides normative constraints on prior probabilities. However, an objective Bayesian conception of IBE must say that H 's explanatory goodness provides normative constraints on priors that positively contribute to its posterior likelihood. That is, the constraints must positively contribute to $P(H)$ or $P(E|H)$, or negatively contribute to $P(E)$. Otherwise, the normative constraints from H 's explanatory goodness will not help us infer that H is true, nor that it is probable.⁵

As we investigate the relationship between IBE and models of scientific explanation, we will, for the most part, ignore the ampliative conception. This is because the ampliative conception can be easily added as an afterthought to almost any model of scientific explanation. Take any model of scientific explanation that provides a set of criteria for what counts as a good explanation. The ampliative explanationist can simply advise that when a hypothesis meets those criteria for being a good explanation, we should give it a probabilistic boost so that its posterior probability is greater than $P(H|E)$. Of course, this doesn't mean models of scientific explanation support the ampliative conception. The models permit giving the best explanatory hypotheses post-conditionalization boosts but do not require it.

For the rest of this paper, my use of 'IBE' excludes the ampliative conception, unless specified otherwise. Our targets are the heuristic conception and the objective Bayesian conception. These are the versions of IBE that can conflict or cohere with models of scientific explanation. Suppose that an overwhelming majority of the models say that H is a good explanation of E only if $P(E|H)$ is high. This would put objective explanationists and most philosophers of science on the same side. On the other hand, perhaps many models of scientific explanation suggest that H 's status as a good potential explanation of E tells us nothing about $P(E|H)$. If so, then objective explanationists and most philosophers of science would be at odds with one another. Similar points hold for

⁵ For example, Huemer ([2009]) suggests using explanatory considerations to help apply the principle of indifference. On their view, we don't infer that H is true or probable based on the fact that it is the best explanation of a body of evidence. This means Huemer's use of explanatory considerations does not constitute IBE, which Huemer seems to concede.

the heuristic conception. Suppose a particular model judges H to be the best explanation of E . However, our best approximation of $P(H|E)$ is that it's somewhere between 0.015 ± 0.005 . In this case, the model suggests that we don't use explanatory language in a way that allows us to infer the best explanations. Thus, the model would be incompatible with the heuristic conception.

To this end, we want to know what models of scientific explanation say about the relationship between H 's status as a good explanation of E and the prior probabilities, $P(H)$, $P(E|H)$, and $P(E)$. That is, whether the models suggest that, all else being equal, if H_1 potentially explains E better than H_2 , then $P(E|H_1) > P(E|H_2)$, or $P(H_1) > P(H_2)$, or $P(E|\sim H_2) > P(E|\sim H_1)$. Note that explanationists do not always agree on which prior probabilities can be affected by explanatory considerations. Some (Okasha [2000]; Poston [2014]; Cabrera [2017]) suggest that explanatory considerations can help ascertain $P(H)$ and $P(E|H)$. Others (Lipton [2004]; Weisberg [2009]; Schupbach and Sprenger [2011]; Henderson [2014]; Climenhaga [2017]) suggest that explanatory considerations can only affect $P(E|H)$. Finally, Cabrera ([2017]) claims that explanatory considerations can affect $P(E|\sim H)$ —and, therefore, $P(E)$.⁶ This is an issue that I will not discuss. When investigating whether a particular model of scientific explanation supports IBE, I will inquire into all the priors— $P(H)$, $P(E|H)$, and $P(E)$. The reader is then free to draw their own conclusions on how well that model supports IBE, based on which view they find most plausible on this controversy about priors.

Finally, note that IBE requires more than a mere positive correlation between good explanations and likely explanations. In addition to positive correlation, some philosophers argue that explanatory considerations must be more epistemically accessible than Bayesian priors (Lipton [2001]; Salmon [2001a], [2001b]; McGrew [2003]). For this reason, Lipton ([2004]) distinguishes between the likeliest explanation—the explanation that is most probable—and the loveliest explanation—the explanation that grants the most understanding. IBE, according to Lipton, moves from the statement that H is the loveliest explanation of E , to the statement that H is the likeliest explanation of E . This distinction helps ensure that IBE involves a non-trivial inference from the explanatory goodness of H to $P(H)$ or $P(E|H)$ is high (or some conclusion like this) and not the other way around. Thus, IBE cannot be ubiquitous in ordinary life and science unless loveliness is often more epistemically accessible than likeliness. If a model of scientific explanation suggests that loveliness is often less epistemically accessible than likelihood, then it is incompatible with IBE.

For precision, I'll borrow Lipton's distinction between loveliness and likelihood, but with a slight modification. In my usage, where M is a model of scientific explanation,

⁶ Cabrera ([2017]) claims that $\sim H$ can be interpreted as a disjunction of all the alternatives to H . So, ascertaining $P(E|\sim H)$ involves the evaluation of how good (or bad) the other competing hypotheses are in comparison to H . Probabilistically, $P(E) = P(E|H)P(H) + P(E|\sim H)P(\sim H)$. So, normative constraints on $P(E|\sim H)$ affect $P(E)$ as well.

H is a lovely potential M explanation of E if and only if H fulfills all the criteria for being a good M explanation/explanans of E, with the possible exception of truth. For example, H is a lovely potential deductive-nomological (DN) explanation of E if and only if H fulfills all the criteria for being a good DN explanation/explanans of E, with the possible exception of truth.

To summarize, a model of scientific explanation supports IBE just in case it entails that (1) explanatory loveliness is positively correlated to likelihood and (2) explanatory loveliness is more epistemically accessible than probability assignments. If a particular model entails that (1) is false or that (2) is false, then it is incompatible with IBE. If a particular model is merely consistent with (1) and (2), then it is compatible with IBE but does not support IBE.

3 An Account that Supports IBE: Kitcher's Unificationist Account

In this section, I argue that Kitcher's unificationist account of explanation supports IBE. If H is a lovely potential unifying explanation of E, $P(E|H)$ is high. Arguably, $P(H)$ is also high, but any general argument we make for the conclusion that $P(H)$ is high will also make $P(E)$ high. The two values, therefore, may offset one another. Furthermore, I argue that unification is a feature of hypotheses that is often more epistemically accessible than prior probabilities.

The unificationist account starts with the intuition that we explain by reducing the number of facts that we have to take as brute. Michael Friedman ([1974]) is the first to develop and defend a unificationist approach to scientific explanation. However, Friedman's unificationist account suffers from formal difficulties (Kitcher [1976]). Consequently, Kitcher's ([1989]) is considered the best version of the unificationist account and it is the version that we will use.

First, we'll introduce the basic concepts Kitcher uses in the unificationist account. A schematic sentence is 'an expression obtained by replacing some, but not necessarily all, the non-logical expressions occurring in a sentence with dummy letters' ([1989], p. 432). A set of filling instructions for a schematic sentence tells us how the dummy letters in the sentence should be replaced. A schematic argument is a set of schematic sentences and a classification for a schematic argument tells us the 'inferential characteristics' of the set—which sentences are premises, which sentences are inferred from which, and so on. A general argument pattern is an ordered triple consisting of 'a schematic argument, a set of sets of filling instructions, and a classification for the schematic argument' (Kitcher [1989], p. 432). We use schematic arguments to unify our body of accepted scientific knowledge, K. A schematic argument unifies K to the extent that it can derive large parts of K when properly filled according to the instructions.

An example might help to illustrate the intuition behind Kitcher's unificationist account. We may interpret Newtonian mechanics as providing argument patterns. The schematic premises can be filled with information about the masses, initial positions,

velocities, or accelerations of various objects. The schematic conclusion can be filled with information about the masses, final positions, velocities, or accelerations of those objects. When these patterns are applied to the planets, we have a particular derivation, say, from a set of initial positions and velocities of the planets to their final positions and velocities. When Newtonian mechanics is applied to falling objects on earth, we have a different derivation. But these different derivations instantiate the same argument pattern provided by Newtonian mechanics. Thus, they are unified—and, therefore, explained—by Newtonian mechanics.

First, let's note that unification seems to be the kind of feature that we can discern apart from probability assignments. We have done exactly that with the example of Newtonian mechanics. We've judged that Newtonian mechanics is a unifying explanation without evaluating its prior probability and the prior probability of the phenomena it unifies. We've noted that we can use Newtonian mechanics to derive various phenomena, but we have not given even an estimate of the conditional probability of the phenomena. So, in terms of making explanatory loveliness more accessible than probability assignments, the unificationist account is on the right track.

Now we want to know whether explanatory loveliness is correlated to likeliness on the unificationist account. Suppose H is a lovely potential unifying explanation of E . That means H fulfills all the criteria for being a unifying explanation of E , with the possible exception of being true. What can we say about $P(H)$, $P(E)$, and $P(E|H)$? Before we answer this, we must determine what H and E correspond to in Kitcher's unificationist account. Note that H and E cannot correspond to the different parts of an argument pattern. If they do, then H would be the ordered triple consisting of the schematic premises, the filling instructions for the schematic premises, and the classification for the schematic premises. E , the unificationist explanandum, would be the ordered triple consisting of the schematic conclusions, the filling instructions for the schematic conclusions, and the classification for the schematic conclusions. However, schematic sentences do not have truth values unless they are tautologous or contradictory. Likewise, filling instructions do not have truth values as they are imperatives. For these reasons, we cannot assign values to $P(H)$ and $P(E)$. $P(E|H)$ is also difficult to ascertain. If the schematic conclusion follows deductively from the schematic premises, then we might make some progress towards ascertaining $P(E|H)$. The filling instructions, however, still present a problem for ascertaining $P(E|H)$. For similar reasons, H cannot be the schematic premises and E the schematic conclusion. Before they are filled, they have no semantic content and, therefore, no truth value.

The other option is to interpret particular derivations, that is, instantiations of argument patterns, as explanations. So, let's suppose H is a set of schematic premises that we have filled properly according to the instructions and E is the corresponding filled schematic conclusion. Now, Kitcher defends the high-probability requirement—the thesis that explanations must have explanans that make their explananda highly probable.

Indeed, Kitcher claims that ‘in a certain sense, all explanation is deductive’ ([1989], p. 448).⁷ Thus, where H is a unifying explanation of E , $P(E|H)$ is high. Furthermore, if H is a set of premises that constitute a good argument for E , then $P(E|H) > P(E)$. This makes Kitcher’s unificationist account suitable for supporting IBE. Kitcher’s account of explanatory goodness entails that a good explanation makes the explanandum probable and is confirmed by the explanandum.

However, this should not be the whole story. What makes the unificationist account distinctive is the intuition that particular derivations, in themselves, have little explanatory value. A particular derivation is explanatory in virtue of instantiating a general argument pattern that unifies many different phenomena. But how should this be interpreted in terms of Bayesian priors? McGrew ([2003]) provides two Bayesian accounts of consilience that are worth considering.⁸ First, we have evidential consilience. On this picture, a hypothesis is consilient to the extent that it confers high probability on a conjunction of evidence, where the individual pieces of evidence are independent apart from the hypothesis. Suppose we have a set of evidence, E_1 through E_n and D_1 through D_n . E_1 through E_n are independent while D_1 through D_n are positively relevant to one another. That is, $P(E_1 \& \dots \& E_n) = P(E_1) \times \dots \times P(E_n)$ and $P(D_1 \& \dots \& D_n) > P(D_1) \times \dots \times P(D_n)$. Suppose that a hypothesis, H , is such that $P(E_1 \& \dots \& E_n|H) = P(D_1 \& \dots \& D_n|H)$ and, for any k , $P(E_k|H) = P(D_k|H)$. In this case, H exhibits evidential consilience with regard to E_1 through E_n , but not with regard to D_1 through D_n .

Evidential consilience supports IBE. Given the assumptions listed above, it follows that

$$\frac{P(E_1 \& \dots \& E_n|H)}{P(E_1 \& \dots \& E_n)} > \frac{P(D_1 \& \dots \& D_n|H)}{P(D_1 \& \dots \& D_n)}$$

Thus, by straightforward application of Bayes’ Theorem, it follows that $P(H|E_1 \& \dots \& E_n) > P(H|D_1 \& \dots \& D_n)$. If H is evidentially consilient with regard to E_1 through E_n but not D_1 through D_n , then E_1 through E_n confirms H to a greater degree than D_1 through D_n confirms H (McGrew [2003], p. 562).

The second account is theoretical consilience. This has to do with a hypothesis’ ability to make different pieces of evidence, E_1 through E_n , positively relevant to one another. In other words, H is theoretically consilient with regard to E_1 through E_n if $P(E_1 \& \dots \& E_n|H) > P(E_1|H) \times \dots \times P(E_n|H)$.

Evidential consilience and unification are very similar concepts. The best unifying explanations allow us to derive large and seemingly unrelated parts of K . Plausibly, we may interpret this to mean that the best unifying explanations make seemingly

⁷ Kitcher admits, however, that nothing in the unificationist account entails the high-probability requirement.

⁸ In an earlier draft of this paper, I addressed only theoretical consilience. I thank an anonymous referee for pointing out that McGrew ([2003]) provides two conceptions of consilience.

independent parts of K probable. This is exactly what we get from evidential consilience. However, evidential consilience and unification are, strictly speaking, distinct. Kitcher insists that all explanations are deductive, which sets unification apart from evidential consilience. Furthermore, unlike evidential consilience, unification does not require that the pieces of evidence unified are independent apart from the unifying hypothesis.

Theoretical consilience is also distinct from unification. Unlike theoretical consilience, unification does not require positive relevance. For example, as mentioned above, Newtonian mechanics unifies the motion of the planets and the behaviour of objects on earth. Let H_g be a large conjunction of the statement that Newton's theory of gravity is correct and statements describing the state of the planets at t_0 and the state of a ball in free-fall near the surface of the earth at t_0 . Let E_1 describe the state of the ball at some later time, t_1 , and E_2 describe the state of the planets at t_1 . On Kitcher's view, H_g unifies E_1 and E_2 , because they fit in the same argument pattern. However, H_g is not theoretically consilient. Since Newton's system is deterministic, $P(E_1 \& E_2 | H_g) = P(E_1 | H_g)P(E_2 | H_g) = 1$, so H_g does not make E_1 and E_2 positively relevant. However, Newtonian mechanics is evidentially consilient in this case since it's plausible that the motion of the planets and the behaviour of objects on earth are independent apart from Newtonian mechanics.

So, unification, evidential consilience, and theoretical consilience are all distinct. However, neither theoretical nor evidential consilience is necessary for confirmation. So, the unificationist account can make IBE rational even though consilience is distinct from unification. Furthermore, even though unification is distinct from evidential consilience, the two concepts are deeply similar. Even if some unifying explanations are not evidentially consilient, the loveliest unifying explanations often are evidentially consilient.

Next, we should also consider the implications of the unificationist account on the priors, $P(H)$ and $P(E)$. If H is a lovely potential unifying explanans of E , then H instantiates schematic premises that, when filled according to the proper instructions, describe part of K . Thus, $P(H)$ would be high to the extent that we believe most of the statements in K are true.⁹

However, there is also a challenge. Insofar as E is also part of an instantiation of an argument pattern that unifies K , it seems that $P(H)$ and $P(E)$ would rise and fall together. If $P(H)$ is high because H fits within a pattern that, when properly filled, describes parts of K , then $P(E)$ should also be high. After all, E also fits within a pattern

⁹ This approach may fit best with realism about science. This does not seem costly as many philosophers who defend IBE do so for the further end of defending realism (Boyd [1983]; Laudan [1984], [1990]; Kvanvig [1994]; Day and Botterill [2008]; Psillos [1996], [2009], [2011]) and philosophers who critique IBE do so for the further end of defending antirealism (van Fraassen [1980], [1989]; Ladyman *et al.* [1997]; Wray [2008]; Khalifa [2010]).

that, when properly filled, describes parts of K. This does not mean that $P(H) = P(E)$, nor does it mean that $P(H)$ and $P(E)$ increase proportionally.¹⁰

So, $P(H)$ and $P(E)$ both increase from being part of a pattern that describes parts of K. However, the issue of which value increases more proportionally is underdetermined. In other words, H's being a lovely potential unifying explanation of E is compatible with $P(H)$ being greater (or increasing more) than $P(E)$, but also with $P(E)$ being greater (or increasing more) than $P(H)$. It's also compatible with $P(H) = P(E)$ and with $P(H)$ increasing to the same extent as $P(E)$. Of course, in a particular context, we may have more information about H and E that may settle the issue.¹¹ But if such a case lends extra support to IBE, it will come from the extra information that allows us to ascertain that $P(H)$ is greater (or increases more) than $P(E)$, not from H's unifying powers. So, Kitcher's unificationist account supports IBE primarily by making $P(E|H)$ high.

In summary, where H is a lovely potential unifying explanation of E, $P(E|H)$ is high and greater than $P(E)$. H may or may not be consilient, but it doesn't need to be consilient to be confirmed by the evidence. We may also claim that $P(H)$ and $P(E)$ are both high but that, in itself, neither helps nor hinders IBE. Additionally, unification seems to be the kind of thing we can ascertain prior to ascertaining Bayesian priors. This makes Kitcher's unificationist account suitable for supporting IBE.

4 Models that are Incompatible with IBE

In this section, I argue that Peter Railton's DNP model, Wesley Salmon's SR model, and Bas van Fraassen's erotetic account are incompatible with IBE. I start with Railton's model because it lies within the familiar Hempelian tradition of interpreting laws as essential to scientific explanation.

4.1 Peter Railton's deductive-nomological-probabilistic (DNP) model

On Hempel's models of scientific explanation, nomicity and expectability are individually necessary components of scientific explanation. In providing a scientific explanation of some event, we show that the event is expectable based on laws. Railton argues, however, that sometimes nomicity and expectability come apart.

Consider the following example, modified from Railton ([1978]). Take a relatively short interval of time between t_0 and t_1 . The probability that a particular uranium-238 (^{238}U) isotope alpha-decayed between t_0 and t_1 is low. However, we believe, based on scientific theory, that spontaneous radioactive decay is governed by laws. On Hempel's view, there is no scientific explanation for why this isotope decayed between t_0 and t_1 . That event may be subsumed under laws, but there is no way to make it expectable. Thus, it is an inexplicable event.

¹⁰ I thank an anonymous referee for this point.

¹¹ For instance, if E is a statement that makes precise claims about multiple independent parameters and H is not a precise statement, then $P(E)$ is likely lower than $P(H)$.

Railton rejects this judgment. On their view, nomicity is necessary for explanation but expectability is not. On the DNP model, explaining the decay of a particular isotope requires providing a deductive argument for the conclusion that it had a certain probability. First, we derive the relevant laws governing the decay of ^{238}U isotopes from theory. Second, we state the laws governing alpha decay. It will specify that all ^{238}U isotopes have probability r to decay between t_0 and t_1 . Third, we have a statement reporting a particular fact, that u is a ^{238}U isotope. These statements constitute the explanans of the DNP explanation. The explanandum consists of two parts: first, the derived conclusion that u had a probability of r to decay between t_0 and t_1 ; second, a ‘parenthetic addendum’ (Railton [1978], p. 214) stating that, by chance, u decayed between t_0 and t_1 . Thus, the DNP explanation for u ’s decay between t_0 and t_1 preserves Hempel’s nomicity requirement, insofar as the DNP explanation subsumes the explanandum under a law. However, it does not preserve Hempel’s expectability requirement, as it does not make the explanandum—specifically the parenthetic addendum—expectable.

Railton provides the following defence for their departure from Hempel’s expectability requirement.

[The DNP explanation] does not explain why the decay *had to* take place, nor does it explain why the decay *could be expected to* take place. And a good thing, too: there is no *had to* or *could be expected to* about the decay to explain—it is not only a chance event, but a very improbable one. [The DNP explanation] does explain why the decay *improbably* took place, which is how it did. (Railton [1978], p. 216)

Notice that these claims are not just about when we have explained an event, but also about the goodness of a potential explanation. Railton does not merely claim that a hypothesis can explain an event without making it probable. In the case of improbable radioactive decay, it is a good-making feature of a purported explanation that it makes the explanandum improbable.

So, suppose that H is a conjunction of the following statements:

- (1) u is a ^{238}U isotope.
- (2) ^{238}U has a half-life of 4.468 billion years.
- (3) All ^{238}U isotopes have probability of r to decay between t_0 and t_1 .

where r is accurately calculated based on (2). Additionally, E is the statement that u decayed between t_0 and t_1 .

H is a lovely potential DNP explanation of E . In other words, H fulfills all the criteria for being a good DNP explanation of E , with the possible exception of truth. This means that if H is false—for example, if u is not a ^{238}U isotope—then it fails to actually DNP-explain E and, if it is true, then it is a lovely actual DNP explanation of E . Indeed,

H is a lovely potential DNP explanation of E even though $P(E|H) = r$, which, given the half-life of ^{238}U , is a low value.

Indeed, if all we know about a hypothesis is that it is a lovely potential DNP explanation, we can infer very little about the Bayesian priors of that hypothesis and the evidence it purports to explain.¹² Thus, the constraints that the DNP model puts on prior probabilities are quite minimal and unsuitable for IBE.

If there are constraints on probability assignments, they would have more to do with the features of particular DNP explanations, rather than the DNP model itself. Consider a case where we know that u is either a ^{238}U isotope or a hydrogen-2 (^2H) isotope. Then, we learn E—that u decayed between t_0 and t_1 . In this case, H is a lovely potential DNP explanans of E. Furthermore, $P(E|H)$ is greater than both $P(E)$ and $P(E|\sim H)$. In this context, we can infer the lovely potential DNP explanation available. However, we can draw the inference only because we have information that goes beyond the fact that H is a lovely DNP explanans of E. Namely, that $P(E|\sim H)$ is much lower than $P(E|H)$.

There will be other contexts where we cannot infer the loveliest potential DNP explanation available. Suppose we start out knowing that u is either a ^{238}U isotope or a recently-discovered isotope of type X. Unfortunately, we don't know the half-life of X isotopes. We only know that they are far less stable than ^{238}U isotopes. In this case, $P(E|H)$ is significantly lower than $P(E|\sim H)$. So, H is disconfirmed by E. More precisely, $P(H|E) < P(H)$. However, H still meets all the requirements for being a lovely potential DNP explanation of E and, if it is true, then it is a lovely actual DNP explanation. In fact, it is the loveliest potential DNP explanation available in that context. Since we don't know much about X, we cannot estimate the probability of u decaying between t_0 and t_1 on the hypothesis that it is an X isotope. Thus, we cannot construct a potential DNP explanation from the hypothesis that u is an X isotope.

The examples multiply. Suppose we learn more about X isotopes and manage to measure their half-life. Now, we can construct a lovely potential DNP explanation from $\sim H$. But this does not make H any less lovely as a potential DNP explanation. The result is a case where H and $\sim H$ are equally lovely potential DNP explanations, but only $\sim H$ can be inferred.

All of this suggests that explanatory loveliness, on the DNP model, is not correlated to likelihood. There are many cases where a lovely potential DNP explanation for a body of evidence is not true, not probable, or is disconfirmed by the evidence. So, we conclude that Railton's DNP model is incompatible with IBE.

4.2 Wesley Salmon's statistical-relevance (SR) model

¹² If an event that has zero probability cannot happen, then we can infer that the DNP model requires $P(E|H)$ to be greater than zero. However, it's not clear that an event with zero probability cannot happen. Suppose u decayed at exactly 12 PM. Plausibly, this event has zero prior probability, yet it happened. Of course, whether an event that has zero probability can happen is an issue we cannot settle in this paper.

Next up is Salmon's ([1971]) statistical-relevance (SR) model. I shall argue that, like the DNP model, it is incompatible with IBE. Again, the underlying intuition behind the SR model starts as a departure from Hempel's models. Richard Jeffrey ([1969]) provides an influential insight to undermine Hempel's requirement that explanations confer high probability on their explananda. Suppose we have a truly stochastic process. Jeffrey claims that as long as we understand the mechanism of the process, as well as the fact that it is stochastic, we understand the improbable outcomes of that process just as much as we understand its probable outcomes.

Again, note that Jeffrey's claim does not merely suggest that there are low-probability explanations. It suggests that low-probability explanations can produce just as much understanding as high-probability explanations. The natural interpretation seems to be that explanations that confer low probability to their explananda can be just as good and lovely as those that confer high probability.¹³

Based on Jeffrey's argument, Salmon ([1971]) develops the statistical-relevance (SR) model as an alternative to Hempel's inductive-statistical model. On Salmon's SR model, it is statistical relevance, rather than high probability, that holds the key to scientific explanation. To provide an SR explanation is to give 'an assemblage of facts statistically relevant to the fact-to-be-explained regardless of the degree of probability that results' (Salmon [1989], p. 67). In this way, Salmon explicitly rejects the Hempelian doctrine that an explanation is an argument, as well as the high-probability requirement.

The explanandum of an SR explanation is a statement of the form, 'this object, a , which belongs to class F , has property G '. The SR explanans for statements of this form cites all and only those properties of objects in class F that are statistically relevant to their having property G . To do this, we use a partition of class F , C_1 to C_k . Relevance is defined in terms of the probability of G in each cell of the partition. A partition is relevant to G just in case no two cells of the partition yield the same probability for G . In other words, $P(G|C_i) \neq P(G|C_j)$, where $i \neq j$. A particular cell is homogeneous with regard to G just in case no relevant partition can be made in that cell. A partition is homogeneous when each cell in it is homogeneous.

Thus, the SR explanation for statements of the form, ' a , which is an F is also G ', includes the initial probability, $P(G|F) = n$; a partition, C_1 to C_k , and the probability of G in each cell; and the statement that a belongs to cell C_j .¹⁴

¹³ Some philosophers also suggest that there is a close relation between explanation and understanding. See, for examples, Lipton ([2001b], [2004]); Strevens ([2013]); and McCain ([2015]).

¹⁴ Salmon also uses the screening-off relation to ensure that no irrelevant factors are included in an SR explanation. If a factor is screened-off, it is rendered irrelevant. The screening-off relation is defined thus: C screens-off A from B , just in case $P(B|A \& C) = P(B|C)$. For example, suppose B is the statement that Jamie did not get pregnant this month, A is the statement that Jamie has been taking birth control pills regularly and C is the statement that Jamie had a hysterectomy last year. In this case, C screens-off A from B , because $P(B|A \& C) = P(B|C)$, but A does not screen-off C from B . Thus, if Jamie had a hysterectomy last year, then the fact that Jamie has been taking birth-control pills is rendered irrelevant for their non-pregnancy.

So, let H be a lovely potential SR explanans of E . What can we deduce about the prior probabilities of H , E , and the conditional probability? Nothing significant for IBE-purposes. Salmon explicitly claims that we have an SR explanation regardless of the resulting probability of the explanandum. In other words, the SR model provides no restrictions on $P(E|H)$. It does not need to be high, nor does it need to be higher than $P(E)$.¹⁵ In fact, Salmon explicitly claims that negatively relevant factors can also have explanatory import in an SR explanation. As long as we have a homogeneous relevant partition of the reference class, we would understand all the members of the reference class equally well.¹⁶ This suggests that not only can H explain E despite $P(E|H)$ being low and lower than $P(E)$, H can explain E just as well as a rival hypothesis that makes E probable.

Salmon insists that actual SR explanations have true explanans. Again, however, the distinction between potential and actual explanations is relevant. If all we know is that H is a lovely potential SR explanans for E , then we cannot infer that H is true, nor that $P(H)$ is high. Indeed, there will be plenty of cases where $P(H)$ is low. After all, a lovely potential SR explanans consists of all and only those facts that are statistically relevant to the explanandum, E . In a scenario where many independent facts are statistically relevant to E , H will consist of a long list of facts. This means $P(H)$ is low because, in general, informativeness is negatively correlated to likelihood.

As an illustration, consider the following example, which is modified from one that Salmon ([1989]) uses. Suppose Jones has a strep infection and recovers quickly from it. Let E be the statement that Jones recovers quickly from a strep infection and let $P(E) = 0.4$. Now, we may form a partition to construct an SR explanation of E . Suppose cell C_1 represents strep patients who are treated with penicillin, which increases the probability of quick recovery to 0.8. Cell C_2 represents strep patients who have underlying medical conditions, which decreases the probability of quick recovery to 0.2.

Now, let's consider two competing SR hypotheses that purport to explain Jones' quick recovery. First, H_1 , which provides information about the partition and states that Jones belongs in C_1 . Second, H_2 , which states that Jones belongs in cell C_2 instead of C_1 , but is otherwise identical to H_1 . Thus, $P(E|H_1) = 0.8$ and $P(E|H_2) = 0.2$. Here is a case where H_1 and H_2 are equally lovely potential SR explanations of Jones' quick recovery. Each of them fulfills all the criteria for being a lovely SR explanation of E , with the possible exception of truth. However, by Bayes' Theorem, we can only infer H_1 . Despite being a lovely potential SR explanation, H_2 is disconfirmed by Jones' recovery, because $P(E|H_2) < P(E)$.

¹⁵ Again, the issue of whether events with zero probability can happen is relevant.

¹⁶ Salmon claims that 'whether the posterior [probability] is higher or lower than the prior [probability] is really beside the point' ([1971], p. 65). The insistence that the posterior probability of the explanandum must be higher than the prior comes from the Hempelian tradition of characterizing explanations as arguments, a tradition that Salmon, through the SR model, explicitly resists.

In summary, accepting the SR model involves denying that a hypothesis that confers high probability to the explanandum is thereby a better explanation than one that confers low probability to the explanandum. In other words, it requires denying that explanatory loveliness is positively correlated to likelihood. This makes the SR model incompatible with IBE.

4.3 Bas van Fraassen's erotetic account

Van Fraassen ([1989]) is among the most well-known critics of IBE. Several of their objections to IBE are still being discussed 30 years after they were published. See, for example, (Kvanvig [1994]; Psillos [1996]; Ladyman *et al.* [1997]; Okasha [2000]; Wray [2008]; Huemer [2009]; Weisberg [2009]; Khalifa [2010]; Douven [2013]; Cabrera [2017]). So, why even consider whether their account of explanation is compatible with IBE?

First, van Fraassen's ([1989]) critique of IBE is aimed specifically at the ampliative conception of IBE. So, van Fraassen ([1980]) may consistently hold an account of explanation that supports the heuristic or objective Bayesian conception of IBE. For this reason and the influence of van Fraassen's account, especially as a representative of pragmatic accounts of explanation, we will see whether van Fraassen's erotetic account fits with IBE.

Van Fraassen's account of explanation starts with the claim that an explanation is an answer to a why-question. A why-question, in turn, consists of several elements. First, a why-question has a topic, P_1 , which is the explanandum—the statement for which we seek an explanation. Second, a why-question has a contrast class, X , which is a set containing the topic, P_1 , and a set of alternatives, P_2 to P_j . A why-question arises under the presupposition that its topic is true and every other member of the contrast class is false. Third, a why-question has a tacit relevance relation, R , which specifies the respect in which an explanation is sought. These three concepts help us understand the erotetic account. Van Fraassen claims that a why-question is a request for an answer, A , which bears the relation R to P_1 and which probabilistically favors P_1 over the other members of the contrast class.

To use a modified version of van Fraassen's example, consider the question,

Q: Why did Adam eat the apple?

On the erotetic account, Q is ambiguous, for it does not specify a contrast class. Possible ways to interpret Q include,

Q₁: Why did Adam eat (rather than leave) the apple?

Q₂: Why did Adam eat the apple (rather than the orange)?

Here, we can see that some answers may be adequate for one of these interpretations, but not the others. That Adam was hungry may be a good answer for Q_1 , but not a good answer for Q_2 .

Once we have a potential answer that bears the right relevance relation to the topic of a why-question, we may evaluate it. On van Fraassen's view, we evaluate an answer, A , to a why-question, Q , on three criteria:

- C1) The likelihood of A , given our background information.
- C2) The extent to which A favors the topic, P_1 , over other members of the contrast class.
- C3) Whether there is an alternative answer that performs better than A on (C1) and (C2).

Thus, if A is highly probable given our background information, it earns high marks on (C1); if A favors P_1 over the other members of the contrast class, it earns high marks on (C2); if there is no alternative answer, B , that is more probable than A and favors P_1 over the contrast class to a greater degree, A earns high marks on (C3).

This set of criteria can be naturally translated in terms of Bayesian priors. Suppose H_1 is an answer to a why-question that has E_1 as its topic. Suppose also that the contrast class has E_2 as the only member other than E_1 . (C1) measures $P(H_1)$ —the higher it is, the lovelier H_1 is as a potential erotetic explanation. (C2) measures the difference between $P(E_1|H_1)$ and $P(E_1)$ in comparison to the difference between $P(E_2|H_1)$ and $P(E_2)$.¹⁷ In probabilistic terms, H_1 favors E_1 over E_2 to the extent that $P(E_1|H_1) - P(E_1) > P(E_2|H_1) - P(E_2)$.¹⁸ (C3) compares H_1 to other potential answers. H_1 performs well on that criterion if we do not have an alternative answer that outperforms it on (C1) and (C2).

We can see how the criteria for evaluating answers cohere with IBE. Recall that, on van Fraassen's view, a why-question arises under the presupposition that exactly one member of the contrast class—the topic—is true. So, let's assume that E_1 and E_2 are mutually exclusive and jointly exhaustive alternatives. It follows that if H_1 favors E_1 over E_2 , then $P(E_1|H_1) > P(E_1)$.¹⁹ By Bayes' theorem, it will also follow that $P(H_1|E_1) > P(H_1)$.

¹⁷ A why-question arises under the presupposition that its topic is true and the other members of the contrast class are false, but that does not mean $P(E_1) = 1$ and $P(E_2) = 0$. Van Fraassen ([1980]) claims that when we evaluate an answer to a why-question, we don't use all of our background knowledge. Among other things, we set aside our knowledge that the topic is true and the other members of the contrast class are false.

¹⁸ Van Fraassen does not give us a precise definition of the favoring relation. However, van Fraassen claims that favoring has to do with how well an answer 'shifts the mass of the probability function toward [the topic]' (van Fraassen [1980], p. 148). To my knowledge, Thomas Grimes ([1987]) is the first to offer the interpretation of the favoring relation that I use here.

¹⁹ Since E_1 and E_2 are mutually exclusive and jointly exhaustive, $P(E_1) + P(E_2) = 1$, and $P(E_1|H_1) + P(E_2|H_1) = 1$. By definition, if H_1 favors E_1 over E_2 , then $P(E_1|H_1) - P(E_1) > P(E_2|H_1) - P(E_2)$. So, $P(E_1|H_1) - P(E_1) > (1 - P(E_1|H_1)) - (1 - P(E_1))$. So, $P(E_1|H_1) - P(E_1) > 1 - P(E_1|H_1) - 1 + P(E_1)$. So, $P(E_1|H_1) - P(E_1) > P(E_1) - P(E_1|H_1)$. So, $2P(E_1|H_1) > 2P(E_1)$. So, $P(E_1|H_1) > P(E_1)$.

Van Fraassen's erotetic account entails that a good explanation raises the probability of the evidence it purports to explain and is thus confirmed by the evidence. Indeed, since the degree to which an answer favors the topic is positively correlated to the goodness of that answer, the erotetic account entails that explanatory loveliness is positively correlated to likelihood. So the erotetic account fits IBE insofar as it makes explanatory loveliness positively correlated to likelihood. However, it is incompatible with IBE on the issue of epistemic accessibility. On the erotetic account, probability assignments are the basis for evaluating explanations. We first ascertain $P(H)$ and whether $P(E|H)$ is greater than $P(E)$. We also evaluate the likelihood of H and how well it favors E in comparison to rival hypotheses. Using these values, we evaluate H with regard to (C1) through (C3) and thereby judge whether H is a lovely potential erotetic explanation of E . Thus, according to van Fraassen, the explanationist is wrong about the order of inference in explanatory reasoning. We do not start from loveliness and infer likelihood. Rather, we start from likelihood and infer loveliness.²⁰

Admittedly, the explanationist would disagree with van Fraassen and argue that explanatory judgments are often more accessible than probability assignments. McGrew's ([2003]) card sharp example lends some plausibility to this suggestion.²¹ Fortunately, our conclusion does not depend on whether explanatory judgments are, in fact, often more accessible than probability assignments. Regardless of whether it is true or false, van Fraassen's erotetic account is incompatible with IBE. Of course, the epistemic accessibility issue is still important and interesting. If McGrew's example is apt, then van Fraassen's erotetic account is inadequate and explanationists would not be troubled by our conclusion that IBE is incompatible with the erotetic account. However, the epistemic accessibility issue cannot be adequately addressed in this paper.

5 A Model that is (Merely) Compatible with IBE: Salmon's Causal-Mechanical (CM) Model

Despite the influence of the SR model, Salmon disowned it in the 1980s in favor of the causal-mechanical (CM) model. In this section, I argue that Salmon's CM model is neutral with regard to IBE. On its own, it is compatible with IBE, but it does not support IBE. More philosophical work is needed if the explanationist wants to use this model to support IBE. At the same time, more philosophical work may reveal that it cannot help the explanationist's project after all.

The CM model uses several notions related to causality. First, there is the notion of a causal process—a process that is capable of transmitting a spatiotemporally continuous mark. When two or more causal processes intersect, they sometimes—but not always—interact. Salmon provides the example of two intersecting rays of light. At the time and place of intersection, the two rays modify one another, but that modification does

²⁰ This is one of Salmon's ([2001a]) objections against IBE.

²¹ I thank an anonymous referee for bringing up McGrew's ([2003]) card sharp example.

not persist beyond the intersection. This intersection, therefore, does not qualify as a causal interaction for Salmon. On the other hand, suppose a ray of light passes through a piece of glass and changes its colour. If the colour change persists beyond the intersection of the glass and the light, then that intersection constitutes a causal interaction (Salmon [1998], p. 72).

Salmon uses these notions to analyse scientific explanations. The explanandum, E, in a CM explanation describes an event in terms of the causal process(es) that it constitutes. The explanans, H, describes the causal processes and interactions leading up to the event described in the explanandum (Salmon [1984], [1998]).

How well does the CM model support IBE? First, note that causal facts, plausibly, are epistemically accessible apart from prior probabilities. They also seem to be decent grounds for inferring probability assignments. I can know, for example, that the introduction of an invasive species into an ecosystem causes an indigenous species to flounder. I can know this even if I don't know the likelihood of the effect, given the cause. Additionally, I can reasonably estimate the likelihood that an indigenous species will flounder given that an invasive species has been introduced and my estimate depends, in part, on my knowledge that the two are causally related. Thus, I think the CM model makes loveliness more epistemically accessible than likelihood. On this issue, the CM model fits with IBE.

However, the CM model does not entail that explanatory loveliness is positively correlated to likelihood. On its own, the CM model does not entail any particular position on whether causes raise the probability of their effects.²² So, it seems that the CM model does not support IBE, but it is also consistent with IBE.

On Salmon's ([1998]) own view, causation does not require positive statistical relevance. That is, causes do not always raise the probability of their effects. Salmon holds this view in opposition to philosophers such as Reichenbach ([1956]) and Suppes ([1970]). The following example might lend intuitive support for Salmon's view. Suppose I have a terminal disease. According to my physician, 50% of patients in similar situations as myself die from the disease within a month. My physician offers me a treatment that heals the disease in 95% of patients, but the treatment kills the other 5% of patients within a month. I decide to take the treatment, and I die from the treatment (suppose that death from the treatment is easily distinguishable from death by the disease). Intuitively, this is a case where taking the treatment caused death, even though it decreased the probability of death.²³

Of course, there are ways to save the positive-relevance requirement for causation in light of cases like these. For example, much depends on how we describe the effect. Taking the treatment decreases the likelihood of death, but increases the likelihood of

²² See Dowe and Noordhof ([2004]); Beebe *et al.* ([2009]); and Hájek and Hitchcock ([2016]) for book-length discussions of probability and causation.

²³ Deborah Rosen is considered the first to present this kind of counterexample to the positive relevance requirement.

treatment-induced death. Ultimately, whether causes always make their effects more likely is an issue that cannot be adequately solved in this paper.²⁴

What we can say, however, is that the CM model is compatible with IBE but does not support IBE. The conjunction of Salmon's view on probabilistic causation and the CM model is incompatible with IBE. There will be many instances where H is a lovely potential CM explanation of E , but $P(E|H)$ is not higher than $P(E)$. However, if the explanationist can defend an account of causation where causes raise the probability of their effects, they can use the CM model to support IBE.

6 Conclusion

I've explored several models of scientific explanation. Of the models I explored, one (the unificationist account) supports IBE, one (the CM model) is compatible with but does not support IBE, and the other three (the DNP model, the SR model, and the erotetic account) are incompatible with IBE. What can we learn from this?

Explanationists may argue that IBE-hostile models of scientific explanation are inadequate precisely because they are incompatible with IBE. They may also point out that some philosophers of science consider IBE to be essential for science. See, for example, (Boyd [1983]; McMullin [1984], [1992]; Achinstein [2001]). Ideally, however, we want explanationists to tell us why the philosophers of science who endorse IBE are right and why those who oppose IBE are wrong.

Alternatively, perhaps explanationists and philosophers of science do not use the term, 'explanation', univocally. Cabrera ([2020]) endorses a view like this, claiming that in the context of IBE, an explanation, roughly, is a hypothesis that ranks highly in terms of the explanatory virtues. This conception may be distinct from the conceptions of explanation used outside the context of IBE.²⁵ But if this is the case, then, instead of inference to the best explanation, we should discuss something like inference to the best unifying explanation or inference to the most explanatorily virtuous explanation. However, this strategy of restricting explanatory inferences may come at the cost of ubiquity, as there will be contexts in which we use a conception of explanation that is incompatible with IBE.

Finally, we may note that only the heuristic conception and the objective Bayesian conception of IBE conflict with various models of explanation. The ampliative conception, on the other hand, is compatible with any model of scientific explanation. Once H has been judged as a good explanation of E according to a particular model, the

²⁴ See Dowe ([2004]) and Hitchcock ([2004]) for several suggestions to handle this kind of counterexample.

²⁵ This suggestion is not unprecedented. Schupbach and Sprenger ([2011]), for example, claim to develop an account of explanatory power that is not shared by philosophers of science.

ampliative explanationist suggests giving H bonus probabilistic points. Thus, the final probability of H will be greater than $P(H|E)$.²⁶

Among explanationists, the ampliative conception is not as popular as the objective Bayesian conception and the heuristic conception. Van Fraassen's ([1989]) argument that ampliative IBE is incompatible with Bayesianism seems to be the main reason for this (McGrew [2003]; Huemer [2009]; Weisberg [2009]; Poston [2014]; Climenhaga [2017]). But why should explanationists want their view to be consistent with Bayesianism? Bayes' theorem, understood as a theorem of the probability calculus, should not be doubted. Bayesianism, however, is the controversial philosophical thesis that one's credence for H, upon learning E, should be equal to $P(H|E)$. It is Bayesianism, rather than Bayes' theorem, that the ampliative explanationist must deny.

So, perhaps explanationists should not be too afraid to deny Bayesianism. Objective Bayesian explanationism and heuristic explanationism may have the virtue of being compatible with Bayesianism, but ampliative explanationism has the virtue of being compatible with different models of scientific explanation. Furthermore, there are promising objections against the Dutch Book argument. Some philosophers argue that the Dutch Book argument for Bayesianism is unsound because you can make betting decisions based on inconsistent credences that do not guarantee loss (Hájek [2008]; Hedden [2013]; Pruss [2020]). Others argue that even if ampliative explanationists are subject to Dutch Book losses, there may be other benefits that could justify ampliative IBE (Douven [2013]; Douven and Schupbach [2015]; Douven and Wenmackers [2017]). If my findings in this paper are accurate, explanationists may have good reason to rally to the ampliative conception.

Acknowledgements

I thank Silvana Handjojo, Alex Pruss, the attendees of Baylor University's philosophy department colloquium, and the anonymous referees from BJPS for helpful comments and suggestions.

Yunus Prasetya

Faculty of Liberal Arts, Universitas Pelita Harapan

Tangerang, Indonesia

yunus.prasetya@gmail.com

²⁶ Admittedly, it might be odd to do this with van Fraassen's erotetic account. We would first judge that A is the most probable answer to a why-question and the one that favors the topic to the greatest extent. On that basis, we evaluate that A is the best explanation for the topic. Then, following the ampliative explanationists' advice, we further boost the probability of A. Though this is not logically incoherent, it seems odd to give a statement a probabilistic boost on the basis that it is the most probable from a group of statements. So, perhaps van Fraassen's erotetic account is the least suitable for pairing with the ampliative conception of IBE.

References

- Achinstein, P. [2001]: *The Book of Evidence*, Oxford: Oxford University Press.
- Beebe, H., Hitchcock, C., and Menzies, P. (eds) [2009]: *The Oxford Handbook of Causation*, Oxford: Oxford University Press.
- Bird, A. [2018]: ‘Inference to the Best Explanation, Bayesianism, and Knowledge’, in K. McCain and T. Poston (eds), *Best Explanations: New Essays on Inference to the Best Explanation*, Oxford: Oxford University Press, pp. 98–121
- Boyd, R. [1983]: ‘On the Current Status of the Issue of Scientific Realism’, *Erkenntnis*, **19**, pp. 45–90.
- Cabrera, F. [2017]: ‘Can there be a Bayesian Explanationism? On the Prospects of a Productive Partnership’, *Synthese*, **194**, pp. 1245–72.
- Cabrera, F. [2020]: ‘Does IBE Require a ‘Model’ of Explanation?’, *British Journal for the Philosophy of Science*, **71**, pp. 727–50.
- Climenhaga, N. [2017]: ‘Inference to the Best Explanation Made Incoherent’, *Journal of Philosophy*, **114**, pp. 251–273.
- Day, M. and Botterill, G. S. [2008]: ‘Contrast, Inference, and Scientific Realism’, *Synthese*, **160**, pp. 249–267.
- Dellsen, F. [2018]: ‘The Heuristic Conception of Inference to the Best Explanation’, *Philosophical Studies*, **175**, pp. 1745–66.
- Douven, I. [2013]: ‘Inference to the Best Explanation, Dutch Books, and Inaccuracy Minimisation’, *The Philosophical Quarterly*, **63**, pp. 428–44.
- Douven, I. and Mirabile, P. [2018]: ‘Best, Second-Best, and Good-Enough Explanations: How They Matter to Reasoning’, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **44**, pp. 1792–813.
- Douven, I. and Schupbach, J. N. [2015]: ‘Probabilistic Alternatives to Bayesianism: The Case of Explanationism’, *Frontiers in Psychology*, **6**, pp. 1–9.
- Douven, I. and Wenmackers, S. [2017]: ‘Inference to the Best Explanation versus Bayes’s Rule in a Social Setting’, *British Journal for the Philosophy of Science*, **68**, pp. 535–570.
- Dowe, P. [2004]: ‘Chance-Lowering Causes’, in P. Dowe and P. Noordhof (eds), *Cause and Chance: Causation in an Indeterministic World*, London and New York: Routledge, pp. 30–7.
- Dowe, P. and Noordhof, P. (eds) [2004]: *Cause and Chance: Causation in an Indeterministic World*. London and New York: Routledge.
- Friedman, M. [1974]: ‘Explanation and Scientific Understanding’, *Journal of Philosophy*, **71**, pp. 5–19.
- Grimes, T. R. [1987]: ‘Explanation and the Poverty of Pragmatics’, *Erkenntnis*, **27**, pp. 79–92.

- Hájek, A. [2008]. Arguments for, or Against, Probabilism?', *British Journal for the Philosophy of Science*, **59**, pp. 793–819.
- Hájek, A. and Hitchcock, C. (eds) [2016]: *The Oxford Handbook of Probability and Philosophy*, Oxford: Oxford University Press.
- Harman, G. [1965]: 'Inference to the Best Explanation', *The Philosophical Review*, **74**, pp. 88–95.
- Hartshorne, C. and Weiss, P. (eds) [1935]: *Collected Papers of Charles Sanders Peirce, Vol V: Pragmatism and Pragmaticism*, Cambridge MA: Harvard University Press.
- Hedden, B. [2013]: 'Incoherence Without Exploitability', *Noûs*, **47**, pp. 482–495.
- Henderson, L. [2014]: 'Bayesianism and Inference to the Best Explanation', *British Journal for the Philosophy of Science*, **65**, pp. 687–715.
- Hitchcock, C. [2004]: 'Do All and Only Causes Raise the Probabilities of Effects?', in J. Collins, N. Hall and L. A. Paul (eds), *Causation and Counterfactuals*, Cambridge MA: MIT Press, pp. 403–18.
- Huemer, M. [2009]: 'Explanationist Aid for the Theory of Inductive Logic', *British Journal for the Philosophy of Science*, **60**, pp. 345–75.
- Jeffrey, R. C. [1969]: 'Statistical Explanation vs. Statistical Inference', in N. Rescher (ed.), *Essays in Honor of Carl G. Hempel*, Dordrecht: D. Reidel, pp. 104–13.
- Khalifa, K. [2010]: 'Default Privilege and Bad Lots: Underconsideration and Explanatory Inference', *International Studies in the Philosophy of Science*, **24**, pp. 91–105.
- Kitcher, P. [1976]: 'Explanation, Conjunction, and Unification', *Journal of Philosophy*, **73**, pp. 207–12.
- Kitcher, P. [1989]. 'Explanatory Unification and the Causal Structure of the World', in P. Kitcher and W. C. Salmon, *Scientific Explanation*, Minneapolis: University of Minnesota Press, pp. 410–505.
- Kvanvig, J. L. [1994]: A Critique of van Fraassen's Voluntaristic Epistemology', *Synthese*, **98**, pp. 325–48.
- Ladyman, J., Douven, I., Horsten, L., and van Fraassen, B. [1997]: A Defence of van Fraassen's Critique of Abductive Inference', *The Philosophical Quarterly*, **47**, pp. 305–21.
- Laudan, L. [1984]: 'Explaining the Success of Science: Beyond Epistemic Realism and Relativism', in J. Cushing, C. F. Delaney, and G. Gutting (eds), *Science and Reality: Recent Work in the Philosophy of Science*, Notre Dame: University of Notre Dame Press, pp. 83–105.
- Laudan, L. [1990]. 'Demystifying Underdetermination', in C. W. Savage (ed.), *Scientific Theories* (Series: Minnesota Studies in the Philosophy of Science, vol. 14), Minneapolis: University of Minnesota Press, pp. 267–97.
- Lipton, P. [2001]: 'What Good is an Explanation?', in G. Hon and S. S. Rakover (eds), *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer, pp. 43–60.

- Lipton, P. [2004]: *Inference to the Best Explanation* (2nd ed.). New York: Routledge.
- Lycan, W. G. [1988]: *Judgment and Justification*, New York: Cambridge University Press.
- McCain, K. [2015]: ‘Explanation and the Nature of Scientific Knowledge’, *Science & Education*, **24**, pp. 827–54.
- McCain, K. and Poston, T. [2014]: ‘Why Explanatoriness is Evidentially Relevant’, *Thought*, **3**, pp. 145–53.
- McCain, K. and Poston, T. [2018]: ‘Best Explanations: An Introduction’, in K. McCain & T. Poston (eds), *Best Explanations: New Essays on Inference to the Best Explanation*, Oxford: Oxford University Press, pp. 2–4.
- McCain, K. and Poston, T. [2019]: ‘Dispelling the Disjunction Objection to Explanatory Inference’, *Philosopher’s Imprint*, **19**, pp. 1–8.
- McGrew, T. [2003]: Confirmation, Heuristics, and Explanatory Reasoning’, *British Journal for the Philosophy of Science*, **54**, pp. 553–67.
- McMullin, E. [1984]: ‘A Case for Scientific Realism’, in J. Leplin (ed.), *Scientific Realism*, Los Angeles: University of California Press, pp. 8–40.
- McMullin, E. [1992]. *The Inference that Makes Science*. Milwaukee: Marquette University Press.
- Okasha, S. [2000]: ‘Van Fraassen’s Critique of Inference to the Best Explanation’, *Studies in History and Philosophy of Science*, **31**, pp. 691–710.
- Poston, T. [2014]: *Reason and Explanation: A Defense of Explanatory Coherentism*, Basingstoke: Palgrave Macmillan.
- Pruss, A. R. [2020]: ‘Avoiding Dutch Books Despite Inconsistent Credences’, *Synthese*. <https://doi.org/10.1007/s11229-020-02786-x>
- Psillos, S. [1996]: ‘On van Fraassen’s Critique of Abductive Reasoning’, *The Philosophical Quarterly*, **46**, pp. 31–47.
- Psillos, S. [2007]: ‘The Fine Structure of Inference to the Best Explanation’, *Philosophy and Phenomenological Research*, **74**, pp. 441–448.
- Psillos, S. [2009]: *Knowing the Structure of Nature: Essays on Realism and Explanation*, London: Palgrave Macmillan.
- Psillos, S. [2011]: ‘The Scope and Limits of the No Miracles Argument’, in D. Dieks, W. Gonzales, S. Hartmann, T. Uebel, and M. Weber (eds), *Explanation, Prediction, and Confirmation*, Berlin: Springer, pp. 23–35.
- Railton, P. [1978]: ‘A Deductive-Nomological Model of Probabilistic Explanation’, *Philosophy of Science*, **45**, pp. 206–26.
- Reichenbach, H. [1956]: *The Direction of Time*, Berkeley: University of California Press.
- Salmon, W. C. [1971]: ‘Statistical Explanation’, in W. C. Salmon, R. C. Jeffrey, and J. G. Greeno, *Statistical Explanation and Statistical Relevance*, Pittsburgh: University of Pittsburgh Press, pp. 29–87.
- Salmon, W. C. [1984]: *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

- Salmon, W. C. [1989]: *Four Decades of Scientific Explanation*, Pittsburgh: University of Pittsburgh Press.
- Salmon, W. C. [1998]: 'Scientific Explanation: Causation and Unification', in W. C. Salmon, *Causality and Explanation*. New York: Oxford University Press.
- Salmon, W. C. [2001a]: 'Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation', in G. Hon and S. S. Rakover (eds), *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer, pp. 59–89.
- Salmon, W. C. [2001b]: 'Reflections of a Bashful Bayesian: A Reply to Peter Lipton', In G. Hon and S. S. Rakover (Eds.), *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer, pp. 121–36.
- Schupbach, J. N. [2018]: 'Inference to the Best Explanation, Cleaned Up and Made Respectable', in K. McCain and T. Poston (eds), *Best Explanations: New Essays on Inference to the Best Explanation*, Oxford: Oxford University Press, pp. 40–62.
- Schupbach, J. N. and Sprenger, J. [2011]: 'The Logic of Explanatory Power', *Philosophy of Science*, **78**, pp. 105–27.
- Strevens, M. [2013]: 'No Understanding Without Explanation', *Studies in History and Philosophy of Science*, **44**, pp. 510–15.
- Suppes, P. [1970]: *A Probabilistic Theory of Causality*, Amsterdam: North Holland Publishing Co.
- Thagard, P. [1978]: 'The Best Explanation: Criteria for Theory Choice', *The Journal of Philosophy*, **75**, pp. 76–92.
- van Fraassen, B. [1980]: *The Scientific Image*. Oxford: Oxford University Press.
- van Fraassen, B. [1989]: *Laws and Symmetry*. Oxford: Oxford University Press.
- Weisberg, J. [2009]: 'Locating IBE in the Bayesian Framework', *Synthese*, **167**, pp. 125–43.
- Woodward, J. [2019]: 'Scientific Explanation', in E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, available at <plato.stanford.edu/archives/win2019/entries/scientific-explanation/>
- Wray, K. B. [2008]. 'The Argument from Underconsideration as Grounds for Anti-Realism: A Defence', *International Studies in the Philosophy of Science*, **22**, pp. 317–26.